

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 617 314**

51 Int. Cl.:

|                    |           |
|--------------------|-----------|
| <b>G10L 21/034</b> | (2013.01) |
| <b>H03G 3/24</b>   | (2006.01) |
| <b>G03G 7/00</b>   | (2006.01) |
| <b>H04B 1/66</b>   | (2006.01) |
| <b>G10L 25/18</b>  | (2013.01) |
| <b>G10L 25/45</b>  | (2013.01) |
| <b>H03G 7/00</b>   | (2006.01) |

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **01.04.2014 PCT/US2014/032578**
- 87 Fecha y número de publicación internacional: **09.10.2014 WO2014165543**
- 96 Fecha de presentación y número de la solicitud europea: **01.04.2014 E 14720877 (1)**
- 97 Fecha y número de publicación de la concesión europea: **04.01.2017 EP 2981963**

54 Título: **Aparato de compresión y método para reducir un ruido de cuantización utilizando una expansión espectral avanzada**

30 Prioridad:

**05.04.2013 US 201361809028 P**  
**12.09.2013 US 201361877167 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**16.06.2017**

73 Titular/es:

**DOLBY LABORATORIES LICENSING CORPORATION (50.0%)**  
**1275 Market Street**  
**San Francisco, CA 94103, US y**  
**DOLBY INTERNATIONAL AB (50.0%)**

72 Inventor/es:

**HEDELIN, PER;**  
**BISWAS, ARIJIT;**  
**SCHUG, MICHAEL y**  
**MELKOTE, VINAY**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 617 314 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Aparato de compresión y método para reducir un ruido de cuantización utilizando una expansión espectral avanzada

## 5 REFERENCIA CRUZADA A SOLICITUDES DE PATENTE RELACIONADAS

Esta solicitud reivindica la prioridad para las solicitudes de patentes provisionales de Estados Unidos números 61/809,028 presentada el 5 de abril de 2013 y 61/877,167, presentada el 12 de septiembre de 2013. Un método para la compresión de datos se da a conocer, a modo de ejemplo, en el documento EP2002429 B1.

10

## CAMPO DE LA INVENCION

Una o más formas de realización se refieren, en general, a un procesamiento de señales de audio y más en particular, para reducir el ruido de codificación en *códecs* de audio utilizando técnicas de compresión/expansión (*companding*).

15

## ANTECEDENTES DE LA INVENCION

Numerosos formatos de sonido digitales de gran aceptación utilizan técnicas de compresión de datos, con pérdidas, que rechazan algunos de los datos para reducir las exigencias de memorización o tasas de transmisión de datos. La solicitud de compresión de datos, con pérdidas, no solamente reducir la fidelidad del contenido fuente (p.ej., contenido de audio), sino que puede introducir también una distorsión notable en la forma de artefactos de compresión. Dentro del contexto de los sistemas de codificación de audio, estos artefactos acústicos se denominan ruido de codificación o ruido de cuantización.

20

25

Los sistemas de audio digitales emplean *códecs* (componentes de codificador-decodificador) para comprimir y descomprimir datos de audio en conformidad con un formato de fichero de audio definido o un formato de audio de soporte de flujo continuo. Los *códecs* ponen en práctica algoritmos que intentan representar la señal de audio con un número mínimo de bits al mismo tiempo que conservan una fidelidad lo más alta posible. Las técnicas de compresión, con pérdidas, suelen utilizarse en el funcionamiento de *códecs* de audio sobre un modelo psico-acústico de percepción auditiva humana. Los formatos de audio suelen implicar el uso de una transformada de dominio de tiempo/frecuencia (p.ej., una transformada de coseno discreta modificada – MDCT) y utilizar efectos de enmascaramiento, tales como enmascaramiento de frecuencias o enmascaramiento temporal de modo que algunos sonidos, incluyendo cualquier ruido de cuantización evidente, se oculte o se enmascare mediante un contenido real.

30

35

La mayoría de los sistemas de codificación de audio están basados en tramas. Dentro de una trama, los *códecs* de audio suelen modelar el ruido de codificación en el dominio frecuencial de modo que se haga menos audible. Varios formatos de audio digitales actuales utilizan tramas de tan largas duraciones que una trama puede contener sonidos de varios niveles o intensidades diferentes. Puesto que el ruido de codificación suele ser estacionario en nivel durante la evolución de una trama, el ruido de codificación puede ser más audible durante partes de baja intensidad de la trama. Dicho efecto puede manifestarse como distorsión pre-eco en donde el silencio (o señal de nivel bajo) que precede a un segmento de alta intensidad se saturan mediante el ruido en la señal de audio decodificada. Dicho efecto puede ser más notable en sonidos transitorios o impulsos procedimientos de instrumentación de percusión, tales como castañuelas u otras fuentes sonoras percusivas agudas. Dicha distorsión suele causarse por el ruido de cuantización introducido en el dominio frecuencial que se extiende a través de la ventana de transformada completa del *códec* en el dominio temporal.

40

45

Las medidas actuales para evitar o minimizar los artefactos pre-eco incluyen el uso de filtros. Dichos filtros, sin embargo, introducen una distorsión de fase y una difuminación temporal. Otra posible solución incluye el uso de más pequeñas ventanas de transformada; sin embargo, este método puede reducir notablemente la resolución de la frecuencia.

50

El contenido descrito en la sección de antecedentes de la invención debe asumirse para la técnica anterior simplemente como un resultado de su mención en dicha sección de antecedentes. De modo similar, un problema citado en la sección de antecedentes o asociado con el contenido de dicha sección no debe asumirse que ha sido previamente reconocido en la técnica anterior. El contenido en la sección de antecedentes de la invención simplemente representa diferentes métodos que en sí mismos y a través de ellos pueden ser también considerados como invenciones.

55

## 60 BREVE SUMARIO DE LAS FORMAS DE REALIZACIÓN

La invención se define por un método, un aparato y un soporte de memorización legible por ordenador, respectivamente, para comprimir o expandir una señal de audio en conformidad con las reivindicaciones 1, 7 y 13 a 15.

65

Las formas de realización se refieren a un método para procesar una señal de audio recibida expandiendo la señal

de audio a una gama dinámica expandida mediante un proceso que incluye dividir la señal de audio recibida en una pluralidad de segmentos temporales utilizando una forma de ventana definida, calcular una ganancia de banda ancha para cada segmento temporal en el dominio frecuencial utilizando una media no basada en la energía de una representación en el dominio frecuencial de la señal de audio, y aplicar el valor de ganancia a cada segmento temporal para obtener la señal de audio expandida. Los valores de ganancia de la ganancia de banda ancha aplicados a cada segmento temporal se seleccionan para tener el efecto de amplificar los segmentos de intensidad relativamente alta y atenuar los segmentos de intensidad relativamente baja. Para este método, la señal de audio recibida comprende una señal de audio original que fue comprimida a partir de una gama dinámica original mediante un proceso de compresión que incluye dividir la señal de audio original en una pluralidad de segmentos temporales utilizando una forma de ventana definida, calcular una ganancia de banda ancha en el dominio frecuencial utilizando una media no basada en la energía de muestras en el dominio frecuencial de la señal de audio inicial, y aplicar la ganancia de banda ancha a la señal de audio original. En el proceso de compresión, los valores de ganancia de la ganancia de banda ancha aplicados a cada segmento temporal se seleccionan para tener el efecto de amplificar segmentos de intensidad relativamente baja y atenuar segmentos de intensidad relativamente alta. El proceso de expansión está configurado para restablecer prácticamente la gama dinámica de la señal de audio inicial, y la ganancia de banda ancha del proceso de expansión puede ser prácticamente la inversa de la ganancia de banda ancha del proceso de compresión.

En un sistema que pone en práctica un método de procesamiento de una señal de audio recibida por un proceso de expansión, un componente de batería de filtros puede utilizarse para analizar la señal de audio para obtener su representación en el dominio frecuencial, y la forma de ventana definida para la segmentación en la pluralidad de segmentos temporales puede ser la misma que el filtro de prototipo para la batería de filtros.

De forma análoga, en un sistema que pone en práctica un método de procesamiento de una señal de audio recibida mediante un proceso de compresión, una componente de batería de filtros puede utilizarse para analizar la señal de audio original para obtener su representación en el dominio frecuencial, y la forma de ventana definida para la segmentación en la pluralidad de segmentos temporales puede ser la misma que el filtro prototipo para la batería de filtros. La batería de filtros, en uno u otro caso, puede ser una de entre una batería QMF o una transformada de Fourier de corta duración. En este sistema, una señal recibida para el proceso de expansión se obtiene después de la modificación de la señal comprimida por un codificador de audio que genera un flujo de bits, y un decodificador que decodifica el flujo de bits. El codificador y el decodificador pueden comprender al menos parte del *códec* de audio basado en la transformada. El sistema puede comprender, además, componentes que procesan la información de control que se recibe por intermedio del flujo de bits y determina un estado de activación del proceso de expansión.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

En los siguientes dibujos las referencias numéricas similares se utilizan para referirse a elementos similares. Aunque las Figuras siguientes ilustran varios ejemplos, las una o más realizaciones no están limitadas a los ejemplos ilustrados en las Figuras.

La Figura 1 ilustra un sistema para comprimir y expandir una señal de audio en un *códec* de audio basado en la transformada, en conformidad con una forma de realización.

La Figura 2A ilustra una señal de audio dividida en una pluralidad de segmentos de corta duración, en conformidad con una forma de realización.

La Figura 2B ilustra la señal de audio representada en la Figura 2A después de la aplicación de una ganancia de banda ancha sobre cada uno de los segmentos de corta duración, en conformidad con una forma de realización.

La Figura 3A es un diagrama de flujo que ilustra un método de compresión de una señal de audio, en conformidad con una forma de realización.

La Figura 3B es un diagrama de flujo que ilustra un método de expansión de una señal de audio, en conformidad con una forma de realización.

La Figura 4 es un diagrama de bloques que ilustra un sistema para comprimir una señal de audio, en conformidad con una forma de realización.

La Figura 5 es un diagrama de bloques que ilustra un sistema para expandir una señal de audio, en conformidad con una forma de realización.

La Figura 6 ilustra la división de una señal de audio en una pluralidad de segmentos de corta duración, en conformidad con una forma de realización.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

Los sistemas y métodos se describen para el uso de técnicas de compresión para conseguir el modelado del ruido temporal de ruido de cuantización en un *códec* de audio. Dichas formas de realización incluyen el uso de un algoritmo de compresión puesto en práctica en el dominio QMF para conseguir un modelado temporal del ruido de cuantización. Los procesos incluyen el control del codificador del nivel de compresión del decodificador deseado y la expansión más allá de las aplicaciones monofónicas a la compresión estéreo y de multicanal.

Aspectos de las una o más formas de realización aquí descritas pueden ponerse en práctica en un sistema de audio que procesa señales de audio para la transmisión a través de la red que incluye uno o más ordenadores o dispositivos de procesamiento que ejecutan instrucciones del software. Cualquiera de las formas de realización descritas puede utilizarse sola o junto con otra en cualquier combinación. Aunque varias formas de realización pueden haberse motivado por diversas deficiencias observadas en la técnica anterior, lo que puede indicarse o aludirse en uno o más lugares en la memoria de especificación, las formas de realización no se refieren necesariamente a cualquiera de estas deficiencias. Dicho de otro modo, diferentes formas de realización pueden referirse a deficiencias distintas que pueden examinarse en la memoria descriptiva. Algunas formas de realización solamente pueden referirse parcialmente a algunas deficiencias o solamente una deficiencia que pueda describirse en la memoria, y algunas formas de realización pueden no relacionarse con ninguna de estas deficiencias.

La Figura 1 ilustra un sistema de compresión para reducir el ruido de cuantización en un sistema de procesamiento de audio basado en un *códec*, en conformidad con una forma de realización. La Figura 1 ilustra un sistema de procesamiento de señal de audio que se basa en un codificador de compresión de *códec* de audio (o "codificador base") 106 y un decodificador (o "decodificador base") 112. El codificador 106 codifica el contenido de audio en un flujo de datos o señal para su transmisión a través de la red 110, en donde se decodifica por el decodificador 112 para su reproducción o cualquier procesamiento adicional. En una forma de realización, el codificador 106 y el decodificador 112 del *códec* realizan un método de compresión con pérdidas para reducir las exigencias de memorización y/o tasa de transmisión de datos de los datos de audio digitales, y dicho *códec* puede realizarse como un MP3, Vorbis, Dolby Digital (AC-3), AAC, o un *códec* similar. El método de compresión con pérdidas del *códec* crea un ruido de codificación que suele ser estacionario en nivel durante la evolución de una trama definida por el *códec*. Dicho ruido de codificación suele ser más audible durante las partes de baja intensidad de una trama. El sistema 100 incluye componentes que reducen el ruido de codificación percibido en sistemas de codificación existentes proporcionando un componente de pre-etapa de compresión 104 antes del codificador base 106 del *códec* y un componente post-etapa de expansión 114 que opera en la salida del decodificador base 112. La componente de compresión 104 está configurada para dividir la señal de entrada de audio original 102 en una pluralidad de segmentos temporales utilizando una forma de ventana definida, calcular y aplicar una ganancia de banda ancha en el dominio frecuencial utilizando una media no basada en la energía de muestras en el dominio frecuencial de la señal de audio inicial, en donde los valores de ganancia aplicados a codificada segmento temporal amplifican los segmentos de intensidad relativamente baja y atenúan los segmentos de intensidad relativamente alta. Esta modificación de la ganancia tiene el efecto de comprimir o reducir notablemente la gama dinámica original de la señal de audio de entrada 102. La señal de audio comprimida se codifica luego en el codificador 106, se transmite a través de la red 110 y se decodifica en el decodificador 112. La señal comprimida decodificada se aplica a la entrada de la componente de expansión 114, que está configurada para realizar la operación inversa de la pre-etapa de compresión 104 aplicando valores de ganancia inversos a cada segmento temporal para expandir la ganancia dinámica de la señal de audio comprimida de nuevo a la gama dinámica de la señal de audio de entrada original 102. De este modo, la señal de salida de audio 116 comprime una señal de audio que tiene la gama dinámica original, con el ruido de codificación eliminado mediante el proceso de compresión post-etapa y pre-etapa.

Según se ilustra en la Figura 1, una componente de compresión o pre-etapa de compresión 104 está configurada para reducir la gama dinámica de la señal de audio 102 aplicada a la entrada del codificador base 106. La señal de audio a la entrada se divide en varios segmentos de corta duración. La magnitud o longitud de cada segmento de corta duración es una fracción del tamaño de la trama utilizada por el codificador base 106. A modo de ejemplo, una magnitud de trama típica del codificador base puede ser del orden de magnitud de 40 a 80 milisegundos. En este caso, cada segmento de corta duración puede ser del orden de magnitud de 1 a 3 milisegundos. La componente de compresión 104 calcula un valor de ganancia de banda ancha adecuado para comprimir la señal de audio a la entrada sobre una base de por segmento. Lo que antecede se consigue modificando los segmentos de corta duración de la señal mediante un valor de ganancia adecuado para cada segmento. Los valores de ganancia relativamente grandes se seleccionan para amplificar segmentos de intensidad relativamente baja y los valores de pequeña ganancia se seleccionan para atenuar los segmentos de intensidad alta.

La Figura 2A ilustra una señal de audio dividida en una pluralidad de segmentos de corta duración, en conformidad con una forma de realización, y la Figura 2B ilustra la misma señal de audio después de la aplicación de la ganancia de banda ancha mediante un componente de compresión. Según se ilustra en la Figura 2A, una señal de audio 202 representa un impulso de sonido o transitorio tal como puede producirse por un instrumento de percusión (p.ej., castañuelas). La señal tiene un pico en amplitud según se ilustra en el trazado representativo de la tensión, V, con respecto al tiempo  $t$ . En general, la amplitud de la señal se relaciona con la energía acústica o intensidad del sonido y representa una medida de la potencia acústica en cualquier punto en el tiempo. Cuando la señal de audio 202 se procesa mediante un *códec* de audio basado en trama, partes de la señal se procesan dentro de las tramas de

transformada (p.ej., MDCT) 204. Los sistemas de audio digitales actuales típicos utilizan tramas de relativamente larga duración, de modo que para los sonidos de impulsos de corta duración o transitorios agudos, una trama única puede incluir sonidos de baja intensidad así como de alta intensidad. De este modo, según se ilustra en la Figura 1, la trama MDCT única 204 incluye la parte de impulsos (pico) de la señal de audio así como una cantidad relativamente grande de señal de baja intensidad antes y después del pico. En una forma de realización, un componente de compresión 104 divide la señal en varios segmentos de corta duración 206, y aplica una ganancia de banda ancha a cada segmento con el fin de comprimir la gama dinámica de la señal 202. El número y magnitud de cada segmento de corta duración puede seleccionarse sobre la base de necesidades de aplicación y limitaciones del sistema. En relación con la magnitud de una trama MDCT individual, el número de segmentos de corta duración puede variar desde 12 a 64 segmentos, y puede comprender normalmente 32 segmentos, pero las formas de realización no están limitadas a este respecto.

La Figura 2B ilustra la señal de audio de la Figura 2A después de la aplicación de una ganancia de banda ancha sobre cada uno de los segmentos de corta duración, en conformidad con una forma de realización. Según se ilustra en la Figura 2B, la señal de audio 212 tiene la misma forma relativa que la señal original 202; sin embargo, la amplitud de los segmentos de baja intensidad se ha aumentado mediante la aplicación de la amplificación de valores de ganancia y la amplitud de los de alta intensidad ha sido atenuada mediante la aplicación de la atenuación de valores de ganancia.

La salida del decodificador base 112 es la señal de audio de entrada con una gama dinámica reducida (p.ej., señal 212) más el ruido de cuantización introducido por el codificador base 106. Este ruido de cuantización presenta un nivel casi uniforme a través del tiempo dentro de cada trama. La componente de expansión 114 actúa sobre la señal decodificada para restablecer la gama dinámica de la señal original. Utiliza la misma resolución de corta duración basada en la magnitud del segmento de corta duración 206 e invierte las ganancias aplicadas en la componente de compresión 104. De este modo, la componente de expansión 114 aplica una pequeña ganancia (atenuación) sobre los segmentos que en la señal original tenían una baja intensidad, y habían sido amplificados por el compresor, y aplica una ganancia de gran magnitud (amplificación) sobre segmentos que en la señal original tenían alta intensidad y habían sido atenuados por el compresor. El ruido de cuantización añadido por el codificador base, que tenía una envolvente temporal uniforme es, de este modo, concurrentemente modelado por la ganancia de post-procesador para seguir aproximadamente la envolvente temporal de la señal original. Este procesamiento hace efectivamente que el ruido de cuantización sea menos audible durante los pasajes con tonos bajos. Aunque el ruido puede amplificarse durante el pasaje de alta intensidad, permanece menos audible debido al efecto de enmascaramiento de la señal intensa del propio contenido de audio.

Según se ilustra en la Figura 2A, el proceso de compresión modifica segmentos discretos de la señal de audio individualmente con valores de ganancia respectivos. En algunos casos, lo que antecede da lugar a discontinuidades a la salida de la componente de compresión que puedan causar problemas en el codificador base 106. De modo análogo, las discontinuidades en la ganancia en la componente de expansión 114 podrían dar lugar a discontinuidades en la envolvente del ruido modelado, lo que podría originar chasquidos audibles en la salida de audio 116. Otra cuestión relacionada con la aplicación de valores de ganancia individuales a segmentos de corta duración de la señal de audio está basada en el hecho de que las señales de audio típicas son una mezcla de numerosas fuentes individuales. Algunas de estas fuentes pueden ser estacionarias en el transcurso del tiempo y algunas pueden ser transitorias. Una señal estacionaria es generalmente constante en sus parámetros estadísticos en el transcurso del tiempo, mientras que las señales transitorias no suelen ser constantes. Habida cuenta de la naturaleza de banda ancha de los transitorios, su huella en dicha mezcla suele ser más visible a las frecuencias más altas. Un cálculo de ganancia que está basado en la energía a corto plazo (RMS) de la señal tiende a polarizarse hacia las frecuencias bajas más intensas y por ello, está dominada por las fuentes estacionarias y presenta poca variación en el transcurso del tiempo. En consecuencia, este método basado en la energía no suele ser efectivo en el modelado del ruido introducido por el codificador base.

En una forma de realización, el sistema 100 calcula y aplica la ganancia en las componentes de compresión y expansión en una batería de filtros con un filtro prototipo corto con el fin de resolver los posibles problemas asociados con la aplicación de valores individuales de la ganancia. La señal a modificarse (la señal original en la componente de compresión 104, y la salida del decodificador base 112 en la componente de expansión 114) se analiza primero por la batería de filtros y la ganancia de banda ancha se aplica directamente en el dominio frecuencial. El efecto correspondiente en el dominio temporal es para suavizar naturalmente la aplicación de la ganancia en conformidad con la forma del filtro prototipo. Lo que antecede resuelve los problemas de las discontinuidades anteriormente descritas. La señal en el dominio frecuencial modificada se convierte luego al dominio temporal mediante una batería de filtros de síntesis correspondiente. Analizando la señal con una batería de filtros se proporciona acceso a su contenido espectral, y permite el cálculo de una ganancia que refuerza preferentemente la contribución debida a las altas frecuencias (o para reforzar la contribución debida a cualquier contenido espectral que sea débil), proporcionando valores de ganancia que no están dominados por las componentes más intensas en la señal. Lo que antecede resuelve el problema asociado con fuentes de audio que comprenden una mezcla de diferentes fuentes, según se describió con anterioridad. En una forma de realización, el sistema calcula la ganancia utilizando una así denominada  $p$ -norma de las magnitudes espectrales, en donde  $p$  suele ser menor que 2 ( $p < 2$ ). Esto permite un mayor énfasis para el contenido espectral débil, en comparación de

cuando se basa en la energía ( $p = 2$ ).

Según se indicó con anterioridad, el sistema incluye un filtro prototipo para suavizar la aplicación de la ganancia. En general, un filtro prototipo es la forma de ventana básica en una batería de filtros, que se modula por las formas de onda sinusoidales para obtener las respuestas de impulsos para los filtros de sub-bandas diferentes en las baterías de filtros. A modo de ejemplo, una transformada de Fourier de corta duración (STFT) es una batería de filtros, y cada línea de frecuencia de esta transformada es una sub-banda de la batería de filtros. La transformada de Fourier de corta duración se pone en práctica multiplicando una señal con una forma de ventana (una ventana de  $N$  muestras), que podría ser rectangular, una forma derivada de Hann, Kaiser-Bessel (KBD) o alguna otra forma. La señal en forma de ventana se somete luego a una operación de la transformada de Fourier discreta (DFT), para obtener la transformada STFT. La forma de ventana en este caso es el filtro prototipo. La transformada DFT está constituida por funciones de base sinusoidal, cada una con una frecuencia distinta. La forma de ventana multiplicada por una función sinusoidal proporciona luego el filtro para la sub-banda correspondiente a esa frecuencia. Puesto que la forma de ventana es la misma a todas las frecuencias, se refiere como siendo un "prototipo".

En una forma de realización, el sistema utiliza una batería QMF (Filtro Modulado en Cuadratura) para la batería de filtros. En una puesta en práctica particular, la batería QMF puede tener una ventana de 64-pt, que forma el prototipo. Esta ventana modulada por las funciones de coseno y seno (correspondientes a 64 frecuencias igualmente espaciadas) forma los filtros de sub-bandas para la batería QMF. Después de cada aplicación de la función QMF, la ventana se desplaza en 64 muestras, es decir, el solapamiento entre segmentos temporales en este caso es  $640 - 64 = 576$  muestras. Sin embargo, aunque la forma de ventana cubre diez segmentos temporales en este caso ( $640 = 10 \cdot 64$ ), el lóbulo principal de la ventana (en donde sus valores muestras son muy significativos) es de aproximadamente 128 muestras de longitud. De este modo, la longitud efectiva de la ventana es todavía relativamente corta.

En una forma de realización, la componente de expansión 114 invierte, en condiciones ideales, las ganancias aplicadas a la componente de compresión 104. Aunque es posible transmitir las ganancias aplicadas por la componente de compresión por intermedio del flujo de bits al decodificador, dicho método consumiría normalmente una tasa binaria importante. En una forma de realización, el sistema 100 estima, en cambio, las ganancias requeridas por la comparación de expansión 114 directamente a partir de la señal disponible a tal respecto, es decir, la salida del decodificador 112, que no requiere efectivamente ningún bit adicional. La batería de filtros en las componentes de compresión y expansión se seleccionan para ser idénticas con el fin de calcular ganancias que sean inversas entre sí. Además, estas baterías de filtros están sincronizadas en el tiempo de modo que cualesquiera retardos efectivos entre la salida de la componente de compresión 104 y la entrada de la componente de expansión 114 sean múltiplos del paso de la batería de filtros. Si el codificador-decodificador base no tuviera pérdidas, y la batería de filtros proporciona una reconstrucción perfecta, las ganancias en las componentes de compresión y expansión serían inversas exactas entre sí, con lo que se permite una reconstrucción exacta de la señal original. En la práctica, sin embargo, la ganancia aplicada por la componente de expansión 114 es solamente una aproximación próxima de la inversa de la ganancia aplicada por la componente de compresión 104.

En una forma de realización, la batería de filtros utilizada en las componentes de compresión y expansión es una batería QMF. En una aplicación de uso típica, una trama de audio podría ser 4096 muestras de longitud con un solapamiento de 2048 con la trama próxima. A la frecuencia de 48 kHz, dicha trama sería de 85.3 milisegundos de longitud. En cambio, una batería de filtros QMF que se utiliza puede tener un paso operativo de 64 muestras (que es de 1.3 ms de longitud), lo que proporciona una resolución temporal fina para las ganancias. Además, la batería QMF tiene un filtro prototipo alisado que tiene una longitud de 640 muestras lo que asegura que la aplicación de ganancia varíe no bruscamente en el transcurso del tiempo. El análisis con esta batería de filtros QMF proporciona una representación en mosaico de tiempo-frecuencia de la señal. Cada intervalo temporal de QMF es igual a un paso operativo y en cada intervalo temporal de QMF existen 64 sub-bandas uniformemente espaciadas. Como alternativa, podrían utilizarse otras baterías de filtros, tales como una transformada de Fourier de corta duración (STFT) y dicha representación en mosaico de tiempo-frecuencia podría obtenerse todavía.

En una forma de realización, la componente de compresión 104 realiza una etapa de pre-procesamiento que pone en escala la entrada del *códec*. Para esta forma de realización,  $S_t(k)$  es una muestra de batería de filtros valiosa compleja en el intervalo temporal  $t$  y el margen de frecuencias  $k$ . La Figura 6 ilustra la división de una señal de audio en varios intervalos temporales para una gama de frecuencias, en conformidad con una forma de realización. Para la forma de realización del diagrama 600, existen 64 contenedores de frecuencias  $k$  y 32 intervalos temporales  $t$  que producen una pluralidad de mosaicos de tiempo-frecuencia según se ilustra (aunque no necesariamente dibujados a escala. Las pre-etapas de compresión ponen a escala la entrada de *códec* para llegar a ser  $S'_t(k) = S_t(k)/g_t$ . En esta ecuación  $g_t = (\bar{S}_t/S_0)^p$  es una media de ranura normalizada.

En la ecuación anterior, la expresión  $\bar{S}_t = \frac{1}{K} \sum_{k=1}^K |S_t(k)|$  es el nivel absoluto medio/1-norma y  $S_0$  es una constante adecuada. Una  $p$ -norma genérica se define en el este contexto como sigue:

$$\bar{S}_t = \left( \frac{1}{K} \sum_{k=1}^K |S_t(k)|^p \right)^{1/p}$$

5 Se ha demostrado que la 1-norma puede proporcionar resultados notablemente mejores que utilizando la energía (rms/2-norma). El valor del término exponencial  $\gamma$  suele estar en la gama de entre 0 y 1 y puede elegirse para ser 1/3. La constante  $S_0$  garantiza valores de ganancia razonables independientes de la plataforma de realización. A modo de ejemplo, puede ser 1 cuando se pone en práctica en una plataforma en donde todos los valores  $S_t(k)$  podrían estar limitados en valor absoluto a 1. Podrían ser potencialmente diferentes en una plataforma en donde  $S_t(k)$  puede tener un valor absoluto máximo diferente. Podría utilizarse también para cerciorarse de que el valor de ganancia medio a través de un conjunto grande de señales sea próximo a 1. Es decir, podría ser un valor de señal intermedio entre un valor de señal máximo y un valor de señal mínimo determinados a partir de un gran cuerpo de contenido.

10 En el proceso post-etapa realizado por la componente de expansión 114, la salida del *códec* se extiende mediante una ganancia inversa aplicada por la componente de compresión 104. Lo que antecede requiere una réplica exacta o casi exacta de la batería de filtros de la componente de compresión. En este caso,  $\tilde{S}_t(k)$  representa una muestra valorada compleja de esta segunda batería de filtros. La componente de expansión 114 pone a escala la salida del *códec* para llegar a ser  $\tilde{S}'_t(k) = \tilde{S}_t(k) \cdot \tilde{g}_t$ .

20 En la ecuación anterior,  $\tilde{g}_t$  es una media de ranura normalizada dada como:

$$\tilde{g}_t = \left( \bar{S}_t / S_0 \right)^{\gamma / (1-\gamma)}$$

25 y

$$\bar{\tilde{S}}_t = \left( \frac{1}{K} \sum_{k=1}^K |\tilde{S}_t(k)|^p \right)^{1/p}$$

30 En general, la componente de expansión 114 utilizará la misma p-norma que se utiliza en la componente de compresión 104. De este modo, si el nivel absoluto medio se utiliza para definir  $\bar{S}_t$  en la componente de compresión 104,  $\bar{\tilde{S}}_t$  se define también utilizando la 1-norma ( $p = 1$ ) en la ecuación anterior.

35 Cuando una batería de filtros compleja (que comprende, a la vez, funciones de base seno y coseno), tal como la transformada STFT o la QMF compleja se utiliza en las componentes de compresión y expansión, el cálculo de la magnitud  $|\tilde{S}_t(k)|$  o  $|S_t(k)|$  de una muestra de sub-banda compleja requiere una operación de raíz cuadrada intensa desde el punto de vista del cálculo. Lo que antecede puede eludirse aproximando la magnitud de la muestra de sub-banda compleja en una diversidad de formas, a modo de ejemplo, sumando la magnitud de sus partes real e imaginaria.

40 En las ecuaciones anteriores, el valor K es igual al número de sub-bandas en la batería de filtros, o un valor más bajo. En general, la p-norma podría calcularse utilizando cualquier subconjunto de las sub-bandas en la batería de filtros. Sin embargo, el mismo subconjunto debe emplearse en, a la vez, el codificador 106 y el decodificador 112. En una forma de realización, las partes de alta frecuencia (p.ej., componentes de audio con frecuencia superior a 6 kHz) de la señal de audio podrían codificarse con una herramienta de expansión espectral avanzada (A-SPX). Además puede ser deseable utilizar solamente la señal con frecuencia superior a 1 kHz (o una frecuencia similar) para servir de guía al modelado del ruido. En tal caso, solamente las sub-bandas en la gama de 1 kHz a 6 kHz pueden utilizarse para calcular la p-norma y en consecuencia, el valor de la ganancia. Además, aunque la ganancia se calcule a partir de un subconjunto de sub-bandas, podría aplicarse también a un subconjunto diferente, y posiblemente mayor, de sub-bandas.

45 50 Según se ilustra en la Figura 1, la función de compresión-expansión, denominada *companding*, para el modelado del ruido de cuantización introducido por el codificador base 106 de un *códec* de audio se realiza con dos componentes separadas 104 y 114 para desempeñar algunas funciones de compresión pre-codificador y funciones de expansión post-decodificador. La Figura 3A es un diagrama de flujo que ilustra un método de compresión de una señal de audio en una componente de compresión de pre-codificador, en conformidad con una forma de realización, y la Figura 3B es un diagrama de flujo que ilustra un método de expansión de una señal de audio en una componente de expansión post-decodificador, en conformidad con una forma de realización.

Según se ilustra en la Figura 3A, el proceso 300 se inicia con la componente de compresión que recibe la señal de

audio de entrada (302). Esta componente divide luego la señal de audio en segmentos de corta duración (304) y comprime la señal de audio a una gama dinámica reducida aplicando valores de ganancia de banda ancha a cada uno de los segmentos de corta duración (306). La componente de compresión realiza, además, algún filtrado prototipo y las componentes de la batería de filtros QMF para reducir o eliminar cualesquiera discontinuidades causadas al aplicar diferentes valores de ganancia a segmentos contiguos, según se describió con anterioridad (308). En algunos casos, tales como los basados en el tipo de contenido de audio o algunas características del contenido de audio, las componentes de compresión y expansión de la señal de audio antes y después de las etapas de codificación/decodificación del *códec* de audio pueden degradar, y no mejorar, la calidad de audio de salida. En tales casos, el proceso de *companding* (compresión-expansión) puede desactivarse o modificarse para hacer retornar los niveles de *companding* (compresión/expansión) diferentes. De este modo, la componente de compresión determina la idoneidad de la función de *companding* y/o el nivel óptimo de *companding* requerido para la entrada de señal específica y el entorno de reproducción de audio, entre otras variables (310). Esta etapa de determinación 310 puede presentarse en cualquier punto práctico del proceso 300, tal como antes de la división de la señal de audio 304 o de la compresión de la señal de audio 306. Si la operación de *companding* se considera adecuada, se aplican las ganancias (306) y el codificador codifica luego la señal para su transmisión al decodificador en conformidad con el formato de datos del *códec* (312). Algunos datos de control de *companding*, tales como datos de activación, datos de sincronización, datos de nivel de *companding* y otros datos de control similares pueden transmitirse como parte del flujo de bits para el procesamiento por la componente de expansión.

La Figura 3B es un diagrama de flujo que ilustra un método de expansión de una señal de audio en una componente de expansión post-decodificador, en conformidad con una forma de realización. Según se ilustra en el proceso 350, la etapa del decodificador del *códec* recibe un flujo de bits que codifica la señal de audio procedente de la etapa del codificador (352). El decodificador decodifica luego la señal codificada en conformidad con el formato de datos de *códec* (353). La componente de expansión procesa luego el flujo de bits y aplica cualesquiera datos de control codificados para desactivar la expansión o modificar los parámetros de expansión sobre la base de los datos de control (354). La componente de expansión divide la señal de audio en segmentos temporales utilizando una forma de ventana adecuada (356). En una forma de realización, los segmentos temporales corresponden a los mismos segmentos temporales utilizados por la compresión de compresión. La componente de expansión calcula luego los valores de ganancia adecuados para cada segmento en el dominio frecuencial (358) y aplica los valores de ganancia a cada segmento temporal para expandir la gama dinámica de la señal de audio de nuevo a la gama dinámica original, o cualquier otra gama dinámica adecuada (360).

#### Control de la función *companding*

Las componentes de compresión y de expansión, que comprenden el denominado *compander* (compresor-expansor) del sistema 100 pueden configurarse para aplicar las etapas de pre-procesamiento y de post-procesamiento solamente en un tiempo determinado durante el procesamiento de la señal de audio, o solamente para algunos tipos de contenido de audio. A modo de ejemplo, la función de *companding* puede presentar ventajas para las señales transitorias musicales y vocales. Sin embargo, para otras señales, tales como señales estacionarias, la función de *companding* puede degradar la calidad de audio. En consecuencia, según se ilustra en la Figura 3A, un mecanismo de control de *companding* (*compresión-expansión*) se proporciona como bloque funcional 310 y los datos de control se transmiten desde la componente de compresión 104 a la componente de expansión 114 para coordinar la operación de *companding*. La forma más simple de dicho mecanismo de control es desactivar la función *companding* para los bloques de muestras de audio en donde la aplicación de la función *companding* está degradando la calidad de audio. En una forma de realización, la decisión de activar/desactivar la función *companding* se detecta en el codificador y se trasmite como un elemento de flujo de bits al decodificador de modo que el compresor y el expansor sean capaces de activarse/desactivarse en el mismo intervalo temporal QMF.

La conmutación entre los dos estados dará lugar normalmente a una discontinuidad en la ganancia aplicada, dando lugar a chasquidos o artefactos de conmutación audibles. Las formas de realización incluyen mecanismos para reducir o eliminar estos artefactos de conmutación. En una primera forma de realización, el sistema permite la desactivación de la función *companding* y solamente en tramas en donde la ganancia está próxima a 1. En este caso, existe solamente una pequeña discontinuidad entre la activación/desactivación de la función *companding*. En una segunda forma de realización, un tercer modo de *companding débil* que está comprendido entre el modo de activación y de desactivación, se aplica en una trama de audio entre las tramas de activación o desactivación, y se señala en el flujo de bits. El modo de *companding débil* efectúa una transición lenta del término exponencial y desde su valor por defecto durante la función *companding* puesta a 0, que es el equivalente de la ausencia de dicha función *companding*. En otra forma de realización, el sistema está configurado para no desactivar simplemente la función *companding* sino más bien para aplicar una ganancia media. En algunos casos, la calidad de audio de las señales estacionarias-tonales puede aumentarse si un factor de ganancia constante se aplica a una trama de audio que se asemeja más a los factores de ganancia de tramas de *companding* activado adyacentes que un factor de ganancia constante de 1.0 en una situación de desactivación de la función *companding*. Dicho factor de ganancia puede calcularse promediando todas las ganancias de *companding* (compresión-expansión) a través de una sola trama. Una trama que contiene una ganancia de *companding* media constante se señala, de este modo, en el flujo de bits.

Aunque las formas de realización se describen en el contexto de un canal de audio monofónico, conviene señalar que en una expansión directa se podrían gestionar múltiples canales repitiendo el método individualmente en cada canal. Sin embargo, las señales de audio que comprenden dos o más canales presentan algunas complejidades adicionales que se resuelven mediante formas de realización del sistema de *companding* ilustrado en la Figura 1. La estrategia de la función *companding* debe depender de la similitud entre canales.

A modo de ejemplo, en el caso de señales transitorias de tipo estéreo se ha observado que la función *companding* independiente de los canales individuales puede dar lugar a artefactos de imágenes audibles. En una forma de realización, el sistema determina un valor de ganancia único para cada segmento temporal a partir de las muestras de sub-bandas de ambos canales y utiliza el mismo valor de ganancia para comprimir/expandir las dos señales. Este método suele ser adecuado siempre que los dos canales tengan señales muy similares, en donde la similitud se define utilizando una correlación cruzada, a modo de ejemplo. Un detector calcula la similitud entre canales y conmuta entre la utilización de la función *companding* individual de los canales o de *companding* conjunta de los canales. Las extensiones a más canales dividirían los canales en grupos de canales utilizando criterios de similitud y aplicarían la función *companding* conjunta sobre los grupos. Esta información de agrupamiento puede transmitirse luego por intermedio del flujo de bits.

#### Puesta en práctica del sistema

La Figura 4 es un diagrama de bloques que ilustra un sistema para comprimir una señal de audio en conjunción con una etapa de codificador de un *códec*, en conformidad con una forma de realización. La Figura 4 ilustra un circuito de hardware o sistema que pone en práctica al menos una parte del método de compresión para utilizar en un sistema basado en *códec* que se ilustra en la Figura 3A. Según se ilustra en el sistema 400, una señal de audio de entrada 401 en el dominio temporal se aplica a la entrada de una batería de filtros QMF 402. Esta batería de filtros realiza una operación de análisis que separa la señal de entrada en múltiples componentes en los que cada filtro de banda de paso incluye una sub-banda de frecuencia de la señal original. La reconstrucción de la señal se realiza en una operación de síntesis efectuada por la batería de filtros QMF 410. En la forma de realización a modo de ejemplo ilustrada en la Figura 4, ambas baterías de filtros de análisis y de síntesis gestionan 64 bandas. El codificador base 412 recibe la señal de audio procedente de la batería de filtros de síntesis 410 y genera un flujo de bits 414 codificando la señal de audio en el formato digital adecuado (p.ej., MP3, AAC, etc.).

El sistema 400 incluye un compresor 406 que aplica valores de ganancia a cada uno de los segmentos de corta duración en los que se ha dividido la señal de audio. Lo que antecede genera una señal de audio de gama dinámica comprimida, tal como se ilustra en la Figura 2B. Una unidad de control de la función *companding* 404 analiza la señal de audio para determinar si, o en qué magnitud debe aplicarse la compresión sobre el tipo de señal (p.ej., vocal) o las características de la señal (p.ej., estacionaria frente a transitoria) u otros parámetros pertinentes. La unidad de control 404 puede incluir un mecanismo de detección para detectar la característica de máximo temporal de la señal de audio. Sobre la base de la característica detectada de la señal de audio y algunos criterios predefinidos, la unidad de control 404 envía señales de control adecuadas al compresor 406 para desactivar la función de compresión o modificar los valores de ganancia aplicados a los segmentos de corta duración.

Además de la función *companding*, numerosas otras herramientas de codificación podrían operar también en el dominio de QMF. Una de dichas herramientas es una A-SPX (expansión espectral avanzada), que se ilustra en el bloque funcional 408 de la Figura 4. A-SPX es una técnica que se utiliza para permitir que frecuencias perceptualmente menos importantes sean codificadas con un sistema de codificación más amplio que el de las frecuencias más importantes. A modo de ejemplo, en una herramienta A-SPX en el extremo decodificador, las muestras de sub-bandas QMF desde la frecuencia más baja pueden replicarse en las frecuencias más altas, y la envolvente espectral en la banda de alta frecuencia se modela luego utilizando la información secundaria transmitida desde el codificador al decodificador.

En un sistema en donde la función *companding* y la técnica A-SPX se realizan en el dominio QMF, en el codificador, los datos de envolvente de A-SPX para las más altas frecuencias pueden extraerse a partir de las muestras de sub-bandas todavía no comprimidas según se ilustra en la Figura 4 y puede aplicarse una compresión solamente a las muestras QMF de más baja frecuencia que corresponden a la gama de frecuencias de la señal codificada por el codificador base 412. En el decodificador, 502 representado en la Figura 5, después del análisis de QMF 504 de la señal decodificada, el proceso de expansión 506 se aplica primero, y la operación de A-SPX reproduce posteriormente las muestras de sub-bandas más altas a partir de la señal expandida en las frecuencias inferiores.

En esta forma de realización a modo de ejemplo, la batería de filtros de síntesis QMF 410 en el codificador y la batería de filtros de análisis QMF en el decodificador 504 introducen conjuntamente un retardo de muestra de  $64 \cdot 64 + 1$  (aproximadamente 9 QMF). El retardo de *códec* base en este ejemplo es 3200 muestras (50 ranuras QMF), por lo que el retardo total es de 59 ranuras. Este retardo se cuantifica introduciendo datos de control en el flujo de bits y utilizándolos como el decodificador, de modo que ambas operaciones del compresor del codificador y del expansor del decodificador están en sincronización.

Como alternativa, en el codificador, pueden aplicarse compresiones sobre el ancho de banda completo de la señal

original. Los datos de envolvente de A-SPX pueden posteriormente extraerse a partir de las muestras de sub-bandas comprimidas. En tal caso, el decodificador, después del análisis de QMF, ejecuta primero la función de la herramienta A-SPX para reconstruir primero la señal comprimida de ancho de banda completo. La etapa de expansión se aplica luego para recuperar la señal con su gama dinámica original.

Otra herramienta operativa que podría utilizarse en el dominio QMF puede ser una herramienta de acoplamiento avanzado (AC) (no ilustrada) en la Figura 4. En un sistema de acoplamiento avanzado, dos canales se codifican como una mezcla descendente monoaural con información espacial paramétrica adicional que puede aplicarse en el dominio QMF en el decodificador para reconstruir una salida en estéreo. Cuando el acoplamiento avanzado AC y la función *companding* se utilizan en conjunción entre sí, la herramienta de AC podría situarse después de la etapa de compresión 406 en el codificador, en cuyo caso, se aplicaría antes de la etapa de expansión 506 en el decodificador. Como alternativa, la información secundaria de AC podría extraerse a partir de una señal estéreo no comprimida, en cuyo caso, la herramienta de AC actuaría después de la etapa de expansión 506 en el decodificador. Un modo de AC híbrido puede soportarse también en el que el control AC se utiliza por encima de una determinada frecuencia y la función estéreo discreta se utiliza por debajo de esta frecuencia o, como alternativa, se utiliza una función de estéreo discreta por encima de determinada frecuencia y se utiliza AC por debajo de esta frecuencia.

Según se ilustra en las Figuras 3A y 3B, el flujo de bits transmitido entre la etapa del codificador y la etapa del decodificador del *códec* incluye algunos datos de control. Dichos datos de control constituyen información secundaria que permite al sistema conmutar entre diferentes modos de la función *companding*. Los datos de control de conmutación (para la activación/desactivación de la función *companding*) junto con potencialmente algunos estados intermedios pueden añadirse en el orden de 1 o 2 bits por canal. Otros datos de control pueden incluir una señal para determinar si todos los canales de una función estéreo discreta o configuración de multicanal utilizarán factores de ganancia de *companding* comunes o si deberían calcularse con independencia para cada canal. Dichos datos solamente pueden requerir un bit extra único por canal. Otros elementos de datos de control similares y sus pesos de ponderación de bits adecuados pueden utilizarse dependiendo de las exigencias y limitaciones operativas del sistema.

#### Mecanismo de detección

En una forma de realización, se incluye un mecanismo de control de la función *companding* como parte de la componente de compresión 104 para proporcionar control de la función *companding* en el dominio QMF. El control de la función *companding* puede configurarse sobre la base de varios factores, tales como un tipo de señal de audio. A modo de ejemplo, en la mayoría de las aplicaciones, la función *companding* debe activarse para señales vocales y señales transitorias o cualesquiera otras señales dentro de la clase de señales de valores máximos temporales. El sistema incluye un mecanismo de detección para detectar el máximo de una señal con el fin de ayudar a generar una señal de control adecuada para la función del denominado *compander*.

En una forma de realización, una medida para el máximo temporal  $TP(k)_{frame}$  se calcula a través del contenedor de frecuencia  $k$  para un *códec* base dado, y se calcula utilizando la fórmula siguiente:

$$TP(k)_{frame} = \frac{\sqrt{\frac{1}{T} \sum_{t=0}^{T-1} S_t(k)^4}}{\sqrt{\frac{1}{T} \sum_{t=0}^{T-1} S_t(k)^2}},$$

En la ecuación anterior,  $S_t(k)$  es la señal de sub-banda y  $T$  es el número de ranuras QMF correspondientes a una trama de codificador base. En una forma de realización a modo de ejemplo, el valor de  $T$  puede ser 32. El valor máximo temporal calculado por banda puede utilizarse para clasificar el contenido acústico en dos categorías generales: señales musicales estacionarias y señales transitorias musicales o señales vocales. Si el valor de  $TP(k)_{frame}$  es menor que un valor definido (p.ej., 1.2), la señal en esa sub-banda de la trama es probablemente una señal musical estacionaria. Si el valor de  $TP(k)_{frame}$  es mayor que este valor, en tal caso, la señal es probable que sea del tipo de señales transitorias musicales o señales vocales. Si el valor es mayor que un valor umbral todavía más alto (p.ej., 1.6), la señal es muy probable que sea una señal transitoria musical pura p.ej., castañuelas. Además, se ha observado que para señales que se producen de forma natural, los valores de máximos obtenidos en diferentes bandas eran más o menos similares, y esta característica podría utilizarse para reducir el número de sub-bandas para las cuales ha de calcularse el valor máximo temporal. Sobre la base de esta observación, el sistema puede poner en práctica una de las dos realizaciones siguientes.

En una primera forma de realización, el detector ejecuta el proceso siguiente. Como una primera etapa, calcula el número de bandas que tienen un máximo temporal mayor que 1.6. Como una segunda etapa, calcula luego la media de los valores máximos temporales de las bandas en donde sea menor que 1.6. Si el número de bandas encontradas en la primera etapa es mayor que 51, o si el valor medio determinado en la segunda etapa es mayor

que 1.45, la señal se determina como siendo una señal transitoria musical y en consecuencia, debería activarse la función *companding*. Dicho detector se desactivará, en la mayor parte del tiempo, para el caso de señales vocales. En algunas formas de realización, las señales vocales se codificarán normalmente mediante un codificador vocal separado, y por ello no suele constituir un problema. Sin embargo, en algunos casos, puede ser deseable activar la función *companding* también para las señales vocales. En este caso, un segundo tipo de detector puede ser preferible.

En una forma de realización, este segundo tipo de detector ejecuta el proceso siguiente. Como una primera etapa, calcula el número de bandas que tienen un máximo temporal superior a 1.2. En una segunda etapa, calcula luego la media de los valores máximos temporales de bandas en donde sea menor que 1.2. A continuación, se aplica la regla siguiente: si el resultado de la primera etapa es mayor que 55: activar la función *companding* y si el resultado de la primera etapa es menor que 15: desactivar la función *companding*; si el resultado de la primera etapa está comprendido entre 15 y 55 y el resultado de la segunda etapa es mayor que 1.16: activar a función *companding*; y si el resultado de la primera etapa es comprendido entre 15 y 55 y el resultado de la segunda etapa es menor que 1.16: desactivar la función *companding*. Conviene señalar que los dos tipos de detectores descritos son solamente dos ejemplos de numerosas soluciones posibles para un algoritmo de detector y pueden utilizarse también otros algoritmos similares o como alternativa.

La función de control de *companding* proporcionada por el elemento 404 de la Figura 4 puede ponerse en práctica en cualquier manera adecuada para permitir que se utilice dicha función *companding* o no se utilice sobre la base de determinados modos operativos. A modo de ejemplo, la función *companding* no se suele utilizar en el canal de LFE (efectos de baja frecuencia) de un sistema acústico envolvente y tampoco se utiliza cuando no existe ninguna funcionalidad de A-SPX (esto es, ninguna QMF) puesta en práctica. En una forma de realización, la función de control de *companding* puede proporcionarse por un programa ejecutado por un circuito o elemento basado en procesador, tales como un elemento de control de *companding* 404. A continuación, se proporciona una sintaxis a modo de ejemplo de un segmento de programa que puede realizar el control de la función *companding*, en conformidad con una forma de realización:

```

Companding_control(nCh)
{
    sync_flag=0;
    if (nCh>1){
        sync_flag
    }
    b_needAvg=0
    ch_count=sync_flag?1:nCh
    for (ch=0; ch<ch_count; ch++){
        b_compand_on[ch]
        if (!b_compand_on[ch]){
            b_needAvg=1;
        }
    }
    if (b_needAvg){
        b_compand_avg;
    }
}

```

Los indicadores de *sync\_flag*, *b\_compand\_on[ch]* y *b\_compand\_avg* o elementos de programa pueden estar en el orden de magnitud de 1 bit de longitud, o cualquier otra longitud dependiente de las limitaciones y exigencias del sistema. Conviene señalar que el código de programa ilustrado anteriormente es simplemente un ejemplo de una forma de poner en práctica una función de control de *companding*, y otros programas o componentes de hardware pueden utilizarse para realizar el control de *companding* en conformidad con algunas formas de realización.

Aunque las formas de realización descritas hasta ahora incluyen el proceso de la función *companding* para reducir el ruido de cuantización introducido por un codificador en un *códec*, debe entenderse que aspectos de dicho proceso de la función *companding* puede aplicarse también en sistemas de procesamiento de señales que no incluyan etapas de codificador y decodificador (*códec*). Además, en el caso de que el proceso de *companding* se utilice en conjunción con un *códec*, el *códec* puede basarse en transformada o no basarse en transformada.

Aspectos operativos de los sistemas aquí descritos pueden ponerse en práctica en un entorno de red de procesamiento de sonido basado en ordenador adecuado para procesar ficheros de audio digitalizados o digitales. Partes del sistema de audio adaptativo pueden incluir una o más redes que comprenden cualquier número deseado de máquinas individuales, incluyendo uno o más enrutadores (no ilustrados) que sirven para memorizar y enrutar los

datos transmitidos entre los ordenadores. Dicha red puede construirse sobre la base de varios protocolos de redes diferentes y puede ser la red Internet, una Red de Área Amplia (WAN), una Red de Área Local (LAN) o cualquiera de sus combinaciones.

5 Uno o más de los componentes, bloques, procesos u otros componentes funcionales pueden ponerse en práctica mediante un programa informático que controla la ejecución de un dispositivo informático basado en procesador del sistema. Conviene señalar que las diversas funciones aquí dadas a conocer pueden describirse utilizando cualquier número de combinaciones de hardware, firmware y/o datos y/o instrucciones materializadas en varios soportes legibles por ordenador o legibles por máquina, en términos de su característica de comportamiento, transferencia de registro, componente lógico y/o otras características. Los soportes legibles por ordenador en los que dichos datos formateados y/o instrucciones pueden materializarse incluyen, sin limitación, soportes de memorización físicos (no transitorios), no volátiles, en varias formas, tales como un soporte de memorización de tipo óptico, magnético o de semiconductores.

10  
15 A no ser que el contexto requiera claramente lo contrario, a través de la descripción y de las reivindicaciones, los términos “comprende”, “que comprende” y similares han de interpretarse en un sentido inclusivo a diferencia de un sentido exclusivo o exhaustivo; es decir, en un sentido de “incluyendo, sin limitación”. Las palabras que utilizan el número singular o plural incluyen también el número plural o singular, respectivamente. Además, los términos “aquí”, “en la presente”, “anterior”, “siguiente” y términos de significado similar se refieren a esta solicitud como un conjunto y no a cualquier parte particular de esta solicitud. Cuando la palabra “o” se utiliza en referencia a una lista de dos o más elementos, dicho término cubre toda la interpretación siguiente de la palabra: cualquiera de los elementos en la lista, la totalidad de los elementos de la lista y cualquier combinación de los elementos en la lista.

20  
25 Aunque una o más formas de realización se han descrito a modo de ejemplo y en términos de las formas de realización específicas, ha de entenderse que una o más puestas en práctica no están limitadas a las formas de realización dadas a conocer. Por el contrario, está previsto que cubran varias modificaciones y disposiciones similares que serían evidentes para los expertos en esta técnica dentro del alcance definido por las reivindicaciones adjuntas.

30

**REIVINDICACIONES**

1. Un método de expansión de una señal de audio que comprende:

5 recibir una señal de audio; y

expandir la señal de audio a una gama dinámica expandida mediante un proceso de expansión que incluye: dividir la señal de audio recibida en una pluralidad de segmentos temporales utilizando una forma de ventana definida, calcular una ganancia de banda ancha para cada segmento temporal en el dominio frecuencial utilizando una media no basada en la energía de una representación en el dominio frecuencial de la señal de audio y aplicar valores individuales de ganancia a cada segmento temporal para obtener la señal de audio de gama dinámica expandida, en donde la aplicación de los valores de ganancia individuales amplifica los segmentos de intensidad relativamente alta y atenúa los segmentos de intensidad relativamente baja.

15 2. El método según la reivindicación 1, en donde los segmentos se solapan.

3. El método según la reivindicación 2, en donde una primera batería de filtros se utiliza para analizar la señal de audio con el fin de obtener la representación en dominio frecuencial y la forma de ventana definida corresponde a un filtro de prototipo para la primera batería de filtros.

20 4. El método según la reivindicación 3, en donde la primera batería de filtros es una de entre una batería de filtros modulados en cuadratura (QMF) o una transformada de Fourier de corta duración.

25 5. El método según la reivindicación 3, en donde la ganancia de banda ancha para cada segmento temporal se calcula utilizando muestras de sub-bandas en un subconjunto de sub-bandas en el segmento temporal respectivo.

6. El método según la reivindicación 5, en donde el subconjunto de sub-bandas corresponde a la gama de frecuencias completa abarcada por la primera batería de filtros.

30 7. Un método de compresión de una señal de audio que comprende:

recibir una señal de audio inicial, y

comprimir la señal de audio inicial con el fin de reducir notablemente una gama dinámica original de la señal de audio inicial por intermedio de un proceso de compresión que comprende dividir la señal de audio inicial en una pluralidad de segmentos utilizando una forma de ventana definida, calcular una ganancia de banda ancha en el dominio frecuencial utilizando una media no basada en la energía de muestras en el dominio frecuencial de la señal de audio inicial, y aplicar valores individuales de ganancia a cada segmento de la pluralidad de segmentos para amplificar los segmentos de intensidad relativamente baja y atenuar los segmentos de intensidad relativamente alta.

40 8. El método según la reivindicación 7, en donde los segmentos están solapados y en donde una primera batería de filtros se utiliza para analizar la señal de audio para obtener la representación en dominio frecuencial y la forma de ventana definida que corresponde a un filtro de prototipo para la primera batería de filtros.

45 9. El método según la reivindicación 8, en donde la primera batería de filtros es una de entre una batería de filtros modulados en cuadratura (QMF) o una transformada de Fourier de corta duración.

10. El método según la reivindicación 8, en donde cada valor individual de ganancia se calcula utilizando muestras de sub-bandas en un subconjunto de sub-bandas en un segmento temporal respectivo.

50 11. El método según la reivindicación 10, en donde el subconjunto de sub-bandas corresponde a la totalidad de la gama de frecuencias cubierta por la primera batería de filtros y en donde la ganancia se aplica en el dominio de la primera batería de filtros.

55 12. El método según la reivindicación 10, en donde la ganancia para cada segmento temporal se deriva de la denominada p-norma de las muestras de sub-bandas en cada segmento temporal en donde p es un número real positivo no igual a dos.

13. Un aparato para comprimir una señal de audio que comprende:

60 una primera interfaz para recibir una señal de audio inicial; y

un compresor para comprimir la señal de audio inicial para reducir notablemente una gama dinámica original de la señal de audio inicial dividiendo la señal de audio inicial en una pluralidad de segmentos utilizando una forma de ventana definida, calcular una ganancia de banda ancha en el dominio frecuencial utilizando una media no basada en la energía de muestras en el dominio frecuencial de la señal de audio inicial, y aplicar valores individuales de

65

ganancia a cada segmento de la pluralidad de segmentos para amplificar segmentos de intensidad relativamente baja y para atenuar segmentos de intensidad relativamente alta.

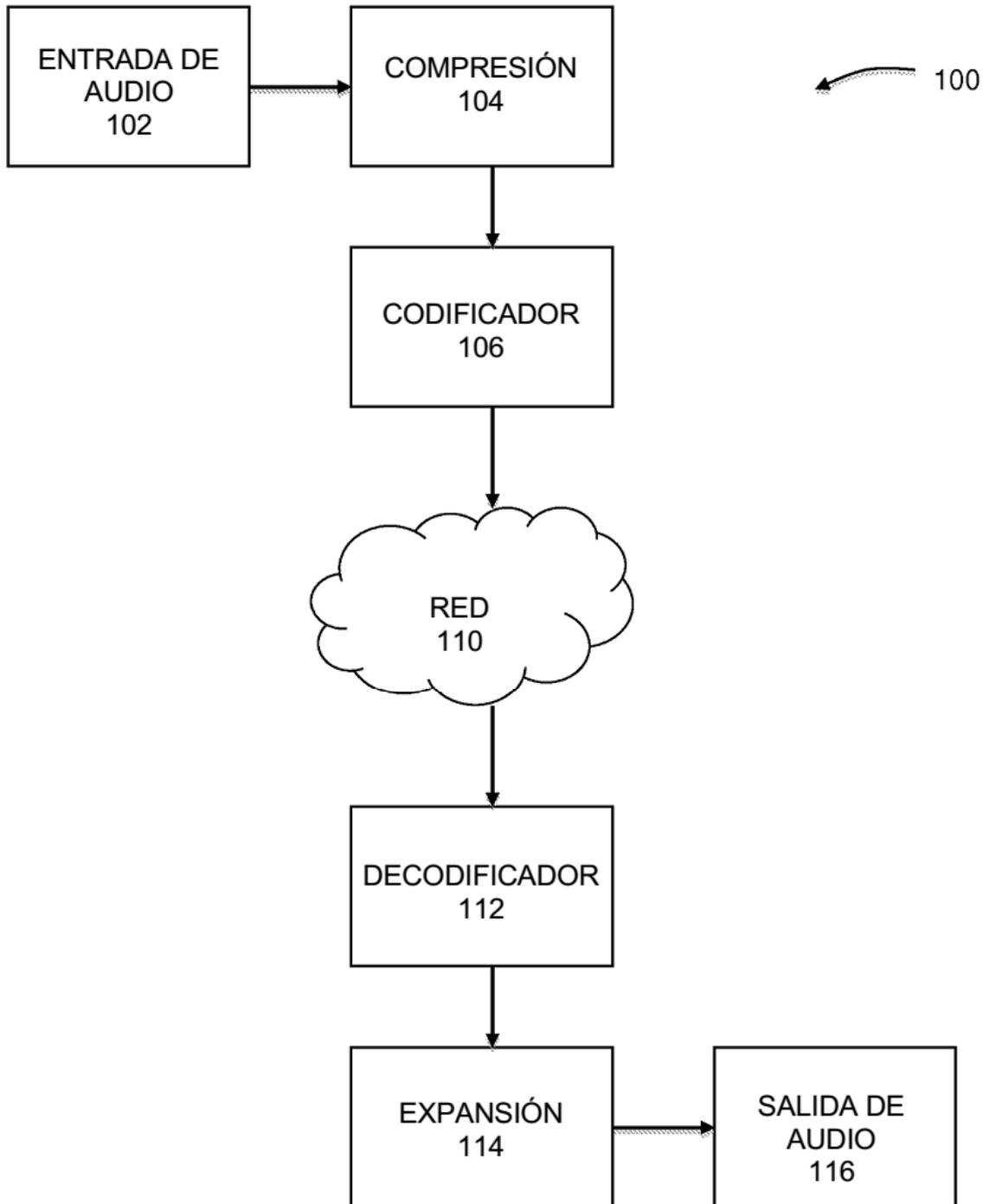
5 **14.** Un aparato para expandir una señal de audio que comprende:

una primera interfaz, para recibir una señal de audio comprimida; y

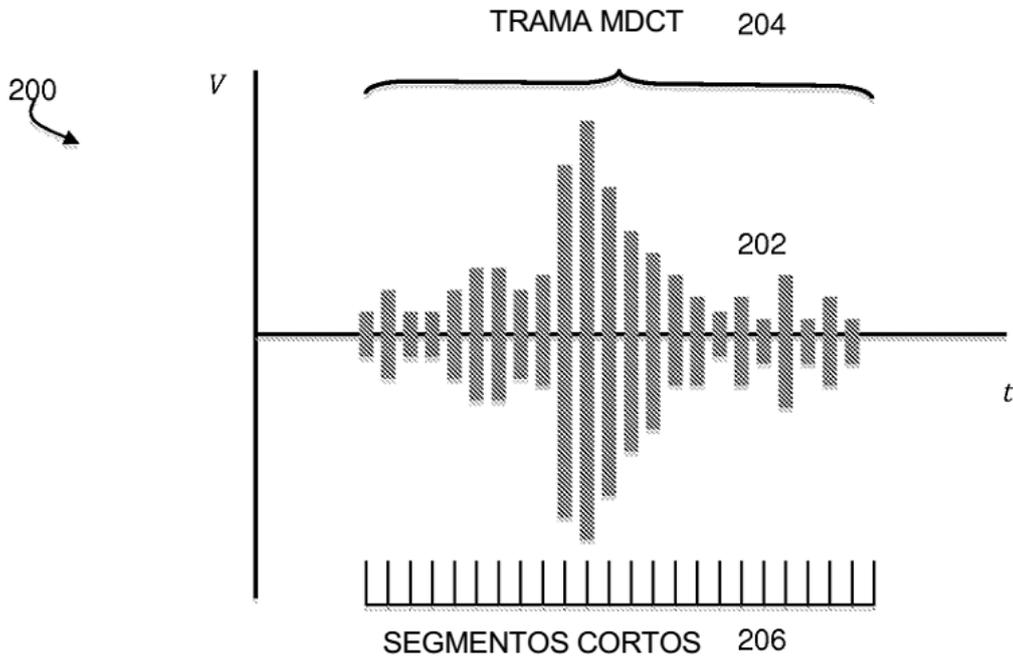
10 un expansor para expandir la señal de audio comprimida para restablecer prácticamente su gama dinámica no comprimida original dividiendo la señal de audio comprimida en una pluralidad de segmentos utilizando una forma de ventana definida, calcular una ganancia de banda ancha en el dominio frecuencial utilizando una media no basada en la energía de muestras en el dominio frecuencial de la señal de audio comprimida, y

15 aplicar valores de ganancia individuales a cada segmento de la pluralidad de segmentos para amplificar segmentos de intensidad relativamente alta y atenuar segmentos de intensidad relativamente baja.

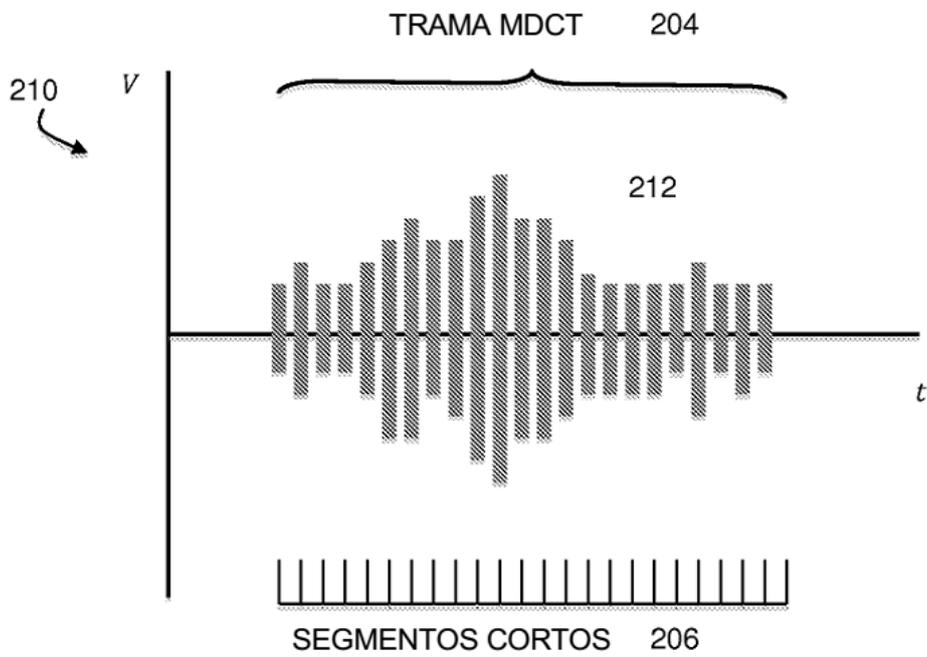
**15.** Un soporte de memorización no transitoria, no volátil, legible por ordenador, que materializa un programa informático con instrucciones adaptadas para realizar el método según cualquiera de las reivindicaciones 1 a 6 o 7 a 12.



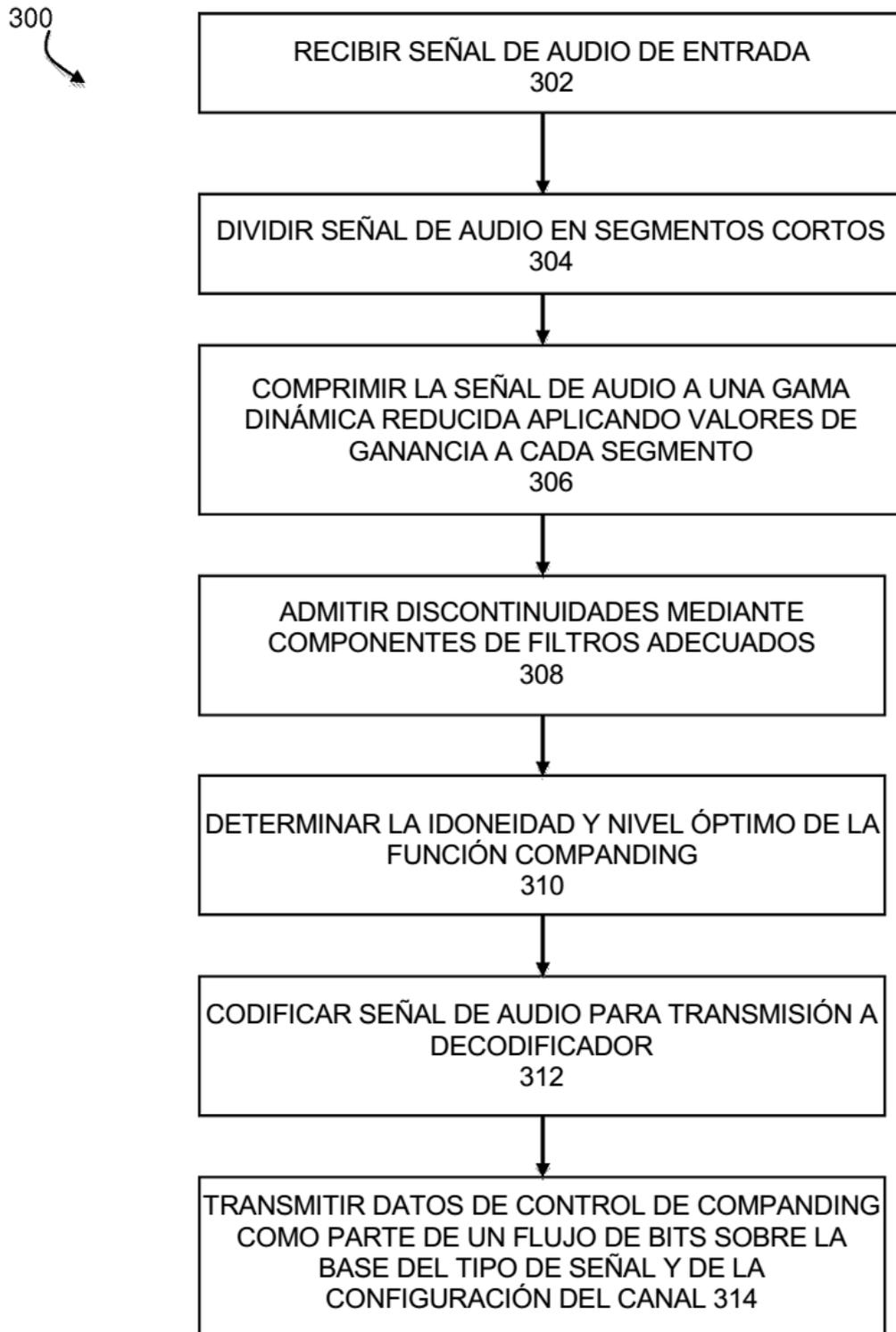
**FIG. 1**



**FIG. 2A**

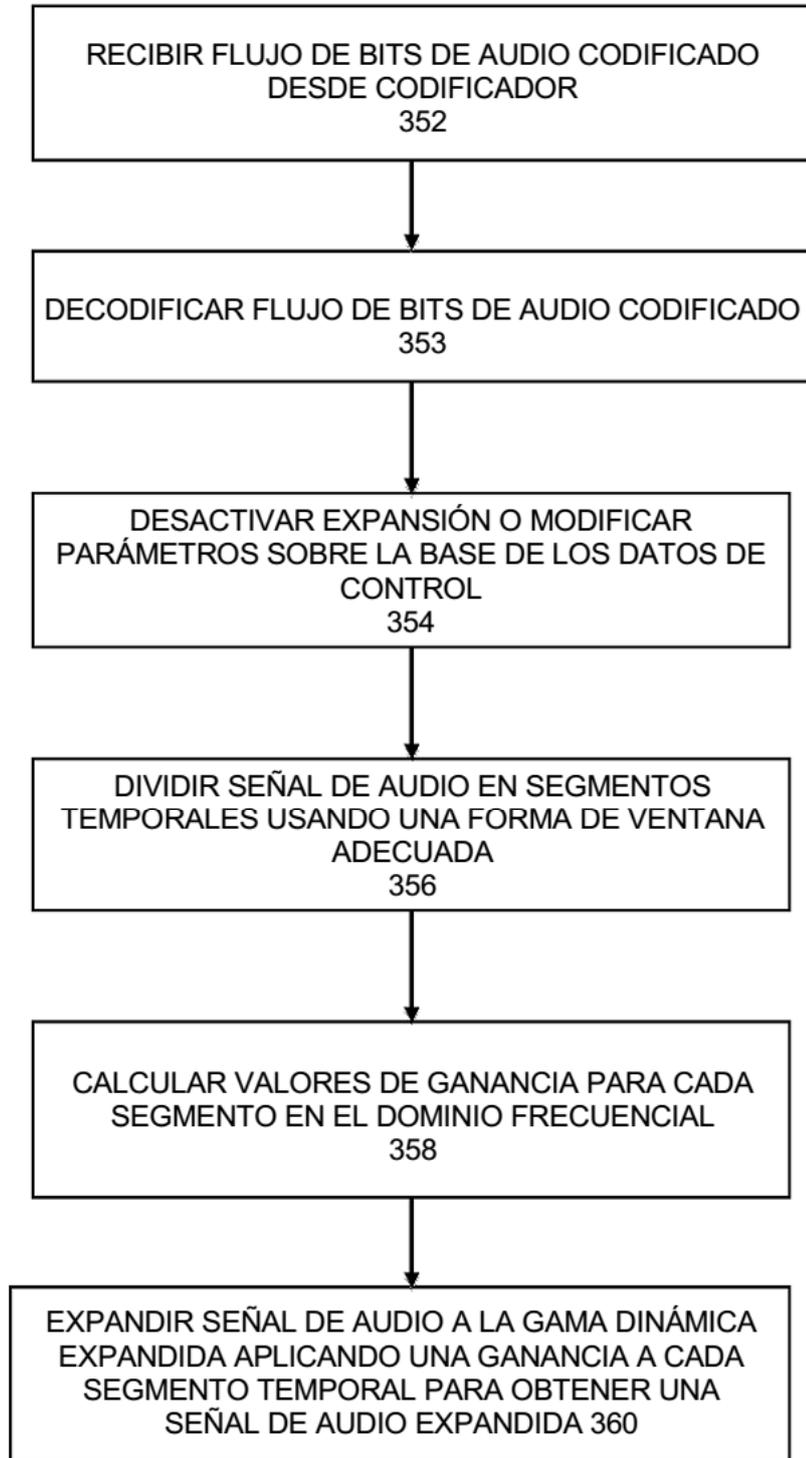


**FIG. 2B**

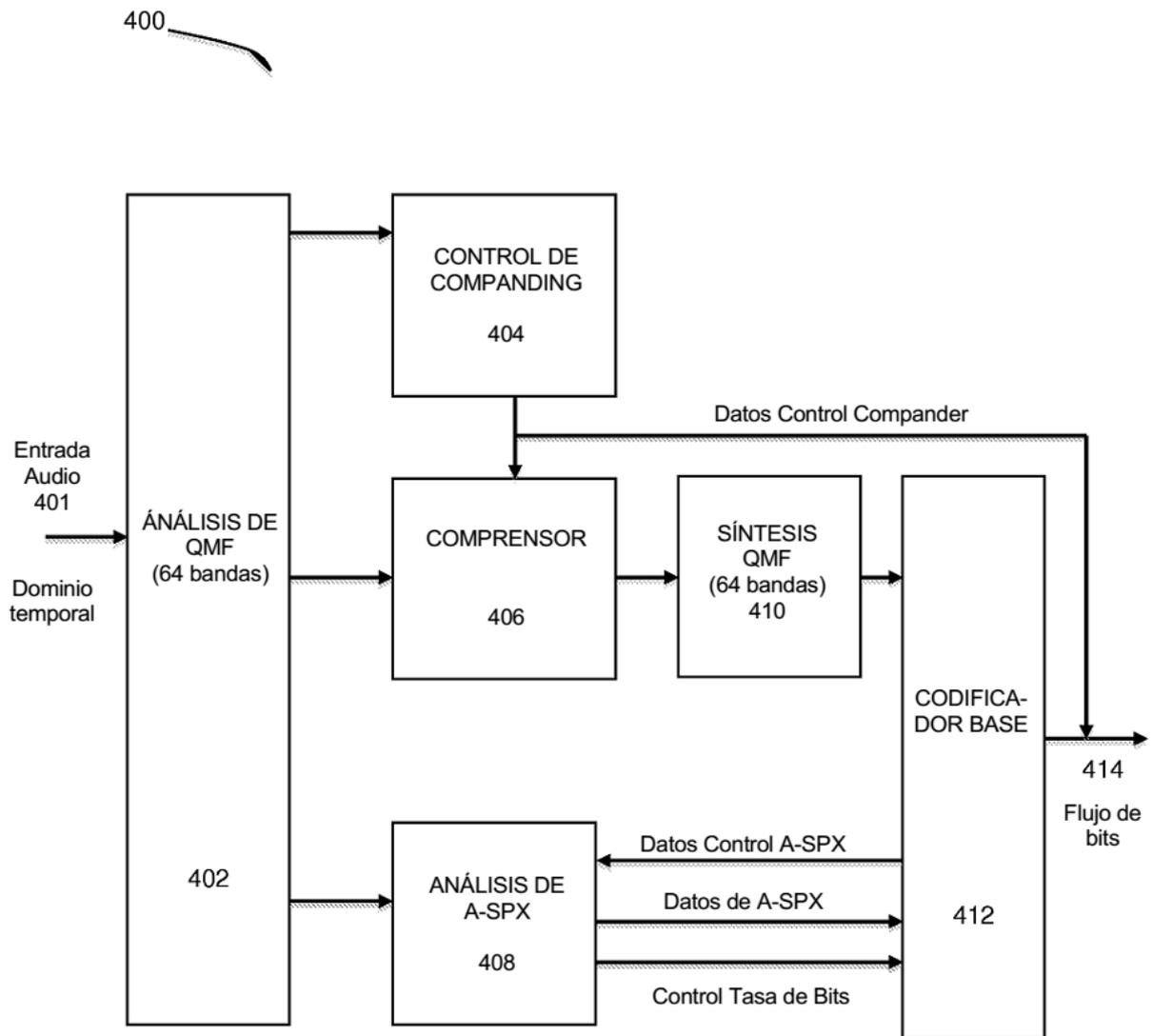


**FIG. 3A**

350  
↘



**FIG. 3B**



**FIG. 4**

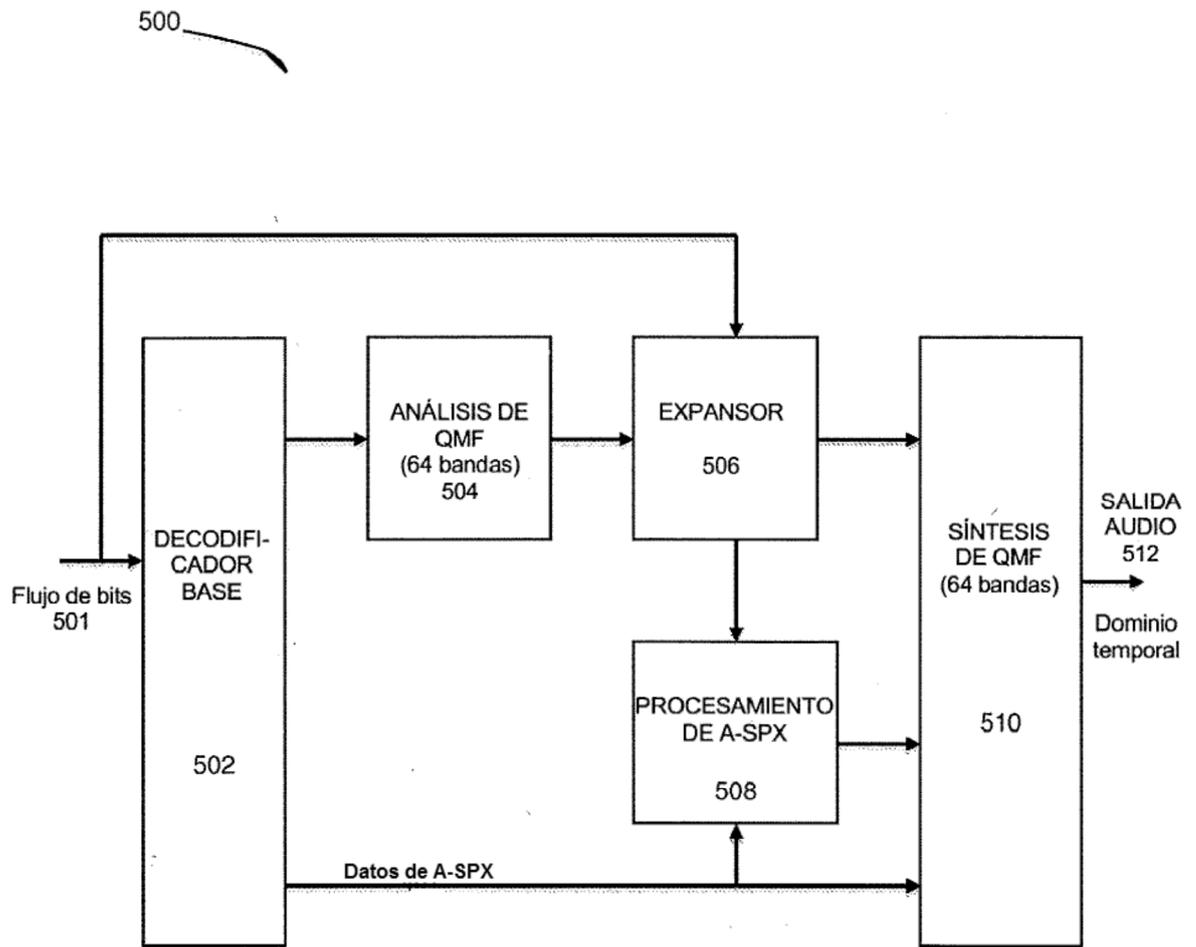
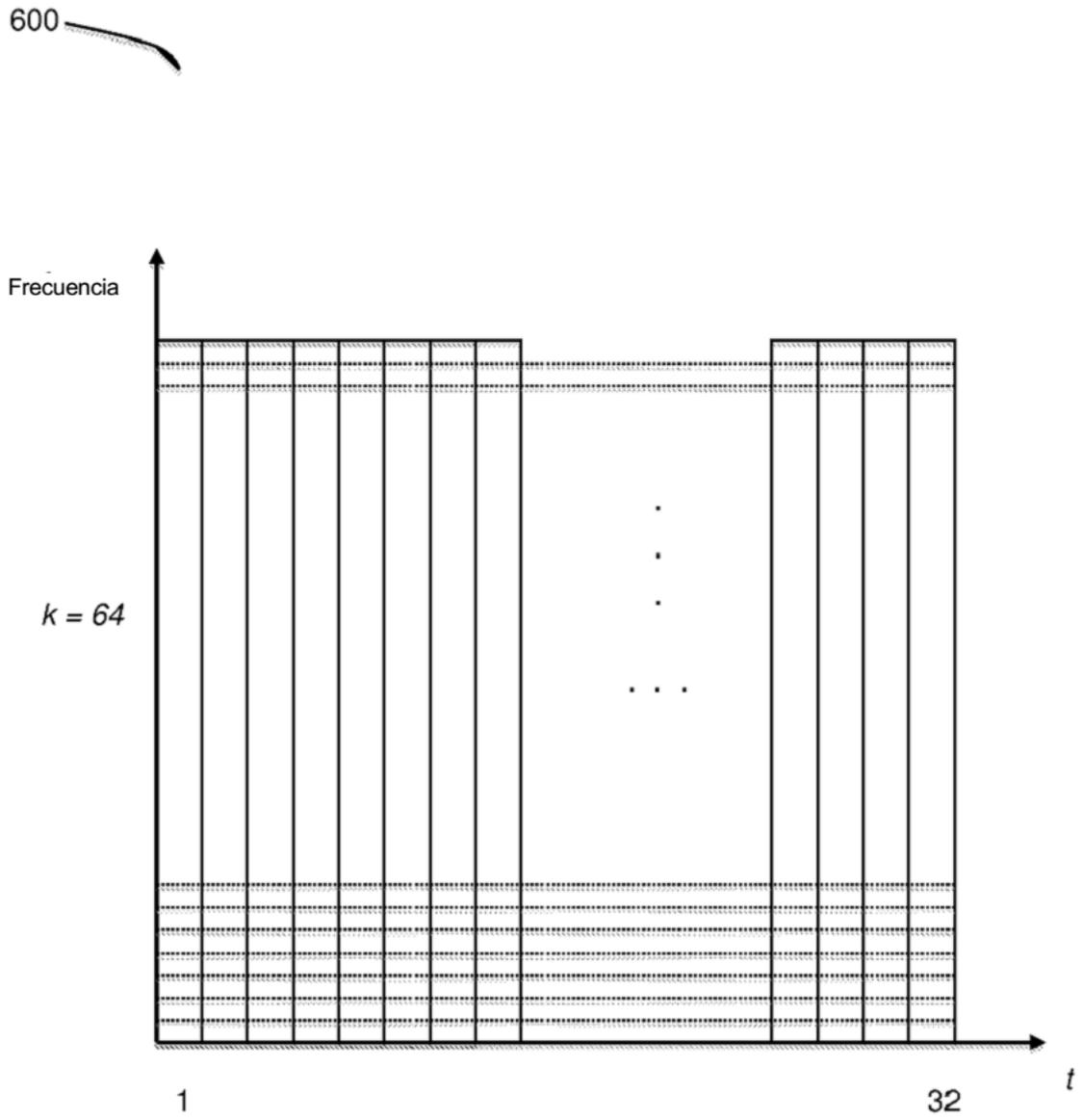


FIG. 5



**FIG. 6**